

EXHIBIT 8



Designation: E 1432 – 04

Standard Practice for Defining and Calculating Individual and Group Sensory Thresholds from Forced-Choice Data Sets of Intermediate Size¹

This standard is issued under the fixed designation E 1432; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ϵ) indicates an editorial change since the last revision or reapproval.

INTRODUCTION

The purpose of this practice is to determine individual sensory thresholds for odor, taste, and other modalities and, when appropriate, calculate group thresholds. The practice takes as its starting point any sensory threshold data set of more than 100 presentations, collected by a forced-choice procedure. The usual procedure is the Three-Alternative Forced-Choice (3-AFC), as exemplified by Dynamic Triangle Olfactometry. A similar practice, Practice E 679, utilizes limited-size data sets of 50 to 100 3-AFC presentations, and is suitable as a rapid method to approximate group thresholds.

Collection of the data is not a part of this practice. The data are assumed to be valid; for example, it is assumed that the stimulus is defined properly, that each subject has been fully trained to recognize the stimulus and did indeed perceive it when it was present above his or her momentary threshold, and that the quality of dilution medium did not vary.

It is recognized that precise threshold values for a given substance do not exist in the same sense that values of vapor pressure exist. A panelist's ability to detect a stimulus varies as a result of random variations in factors such as alertness, attention, fatigue, events at the molecular level, health status, etc., the effects of which can usually be described in terms of a probability function. At low concentrations of an odorant or tastant, the probability of detection by a given individual is typically 0.0 and at high concentrations it is 1.0, and there is a range of concentrations in which the probability of detection is between these limits. By definition, the threshold is the concentration for which the probability of detection of the stimulus is 0.5 (that is, 50 % above chance, by a given individual, under the conditions of the test).

Thresholds may be determined (1) for an individual (or for individuals one by one), and (2) for a group (panel). While the determination of an individual threshold is a definable task, careful consideration of the composition of the group is necessary to ensure the determined threshold represents the group of interest.

There is a large degree of random error associated with estimating the probability of detection from less than approximately 500 3-AFC presentations. The reliability of the results can be increased greatly by enlarging the panel and by replicating the tests.

1. Scope

1.1 The definitions and procedures of this practice apply to the calculation of individual thresholds for any stimulus in any medium, from data sets of intermediate size, that is, consisting

of more than 20 to 40 3-AFC presentations per individual. A group threshold may be calculated using 5 to 15 individual thresholds.

1.2 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

¹ This practice is under the jurisdiction of ASTM Committee E18 on Sensory Evaluation of Materials and Products and is the direct responsibility of Subcommittee E18.04 on Fundamentals of Sensory.

Current edition approved April 1, 2004. Published April 2004. Originally approved in 1991. Last previous edition approved in 1997 as E 1432-91 (1997).

2. Princ

2.1 Th
procedur
measure
triangula

2.2 Fo

practice
scale ste
from the
typically
concentr
ensure th
near the
concentr
typically

2.3 In

for calcu
composit
more) is
threshold
(see 7.2)

2.4 Po

a single s

3. Refer

3.1 AS

E 122

With

terist

E 679

olds

Meth

3.2 C

EN 13

trati

3.3 IS

ISO 1

guid

tion

(3-A

4. Termi

4.1 De

4.1.1

procedur

For exam

etry, the

of which

contain o

containin

² For ref

contact AST

Standards v

the ASTM v

³ Availab

London W4

⁴ Availab

4th Floor, N



2. Principles

2.1 The 3-AFC procedure is one of the set of n-AFC procedures, any of which could be used, in principle, for the measurement of sensory thresholds, as could the duo-trio, the triangular, and the two-out-of-five procedures.

2.2 For calculation of the threshold of one individual, this practice requires data sets taken at five or more concentration scale steps, typically six or seven steps, with each step differing from the previous step by a factor usually between 2 and 4, typically 3.0. The practice presupposes that the range of concentrations has been selected by pretesting, in order to ensure that the individual's threshold falls neither outside nor near the ends of the range, but well within it. At each concentration step, the individual must be tested several times, typically five or more times.

2.3 Individual thresholds, as determined in 2.2, may be used for calculation of a group (or panel) threshold. The size and composition of the panel (usually 5 to 15 members, preferably more) is determined according to the purpose for which the threshold is required and the limitations of the testing situation (see 7.2).

2.4 Pooling of the data sets from panel members to produce a single step calculation of the panel threshold is not permitted.

3. Referenced Documents

3.1 ASTM Standards:²

E 122 Practice for Calculating Sample Size to Estimate, With a Specified Tolerable Error, the Average for Characteristic of a Lot or Process

E 679 Practice for Determination of Odor and Taste Thresholds By a Forced-Choice Ascending Concentration Series Method of Limits

3.2 CEN Standard:³

EN 13725 Air Quality—Determination of Odour Concentration Using Dynamic Dilution Olfactometry

3.3 ISO Standard:⁴

ISO 13301 Sensory Analysis—Methodology—General guidance for Measuring Odour, Flavour, and Taste Detection Thresholds by a Three Alternative Forced Choice (3-AFC) Procedure

4. Terminology

4.1 Definitions of Terms Specific to This Standard:

4.1.1 *Three-Alternative Forced-Choice (3-AFC) test procedure*—a test presentation used in many threshold tests. For example, in odor testing by Dynamic Triangle Olfactometry, the panelist is presented with three gas streams, only one of which contains the diluted odorant, while the other two contain odorless carrier gas. The panelist must indicate the one containing the added substance. (The 3-AFC procedure is

different from the classical Triangle test, in which either one or two of the three samples may contain the added substance.)

4.1.2 *model*—an abstract or concrete analogy, usually mathematical, which represents in a useful way the functional elements of a system or process. In short, the experimenter's theory of what is guiding the results observed.

4.1.3 *statistical model*—a model assuming that the principal factor causing the results to deviate from the true value is a random error process. This can usually be described in terms of a probability function, for example, a bell-shaped curve, symmetrical or skewed. Errors are binomially distributed in the 3-AFC test procedure.

4.1.4 *threshold, detection*—the intensity of the stimulus that has a probability of 0.5 of being detected under the conditions of the test. The probability of detection at any intensity is not a fixed attribute of the observer, but rather a value which assumes that sensitivity varies as a result of random fluctuation in factors such as alertness, attention, fatigue, and events at the molecular level, the effects of which can be modeled by a probability function.

4.1.5 *individual threshold*—a threshold based on a series of judgments by a single panelist.

4.1.6 *group threshold*—the average, median, geometric mean or other agreed measure (or an experimentally determined measure) of central tendency of the individual thresholds of the members of a group (panel). The meaning and significance of the term depends on what the group is selected to represent (see 7.2.2).

4.1.7 *scale step factor*—for a scale of dilutions presented to a panel, the factor by which each step differs from adjacent steps.

4.1.8 *dilution factor*—the following applies to flow olfactometry: If F_1 represents the flow of odorless gas which serves to dilute the flow of odorant, F_2 , the dilution factor, Z , is given by:

$$Z = \frac{F_1 + F_2}{F_2} \quad (1)$$

where Z is dimensionless. F_1 and F_2 may be expressed, both in units of mass, or (preferably) both in units of volume; the report should state which. The term Z_{50} represents the dilution factor to threshold. Alternate terminology in use is as follows: dilution-to-threshold ratio (D/T or D-T); odor unit (OU); and effective dose (ED).

5. Summary of Practice

5.1 From a data set according to 2.2, calculate the threshold for one individual graphically or by linear regression according to 5.2, or by using a model fitting computer program according to 5.3.

5.2 Obtain the threshold in 5.1 by first calculating the proportion correct above chance for each concentration step. This is accomplished by deducting, from the proportion of correct choices, the proportion that would have been selected by chance in the absence of the stimulus (see 8.1.2). Then, for each individual calculate that concentration which has a probability of 0.5 of being detected under the conditions of the test. This is the individual threshold.

² For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

³ Available from British Standards Institute (BSI), 389 Chiswick High Rd., London W4 4AL, U.K.

⁴ Available from American National Standards Institute (ANSI), 25 W. 43rd St., 4th Floor, New York, NY 10036.



5.3 Alternatively obtain the threshold in 5.1 directly from the proportion of correct choices by non-linear regression using a computer program, as described in 8.2.2.

5.4 Always report the individual thresholds of the panelists. Depending on the purpose for which a threshold is required (see 7.2), and on the distribution found, a group threshold may be calculated as the arithmetic or geometric mean, the median, or another measure of central tendency, or it may be concluded that no group threshold can be calculated (see 7.4).

6. Significance and Use

6.1 Sensory thresholds are used to determine the potential of substances at low concentrations to impart odor, taste, skinfeel, etc. to some form of matter.

6.2 Thresholds are used, for example, in setting limits in air pollution, in noise abatement, in water treatment, and in food systems.

6.3 Thresholds are used to characterize and compare the sensitivity of individuals or groups to given stimuli, for example, in medicine, ethnic studies, and the study of animal species.

7. Panel Size and Composition Versus Purpose of Test

7.1 *Panel Size and Composition*—Panel variables should be chosen as a function of the purpose for which the resulting threshold is needed. The important panel variables are as follows:

- 7.1.1 Number of tests per panelist,
- 7.1.2 Number of panelists,
- 7.1.3 Selection of panelists to represent a given population, and
- 7.1.4 Degree of training.

7.2 *Purpose of Test*—It is useful to distinguish the following three categories:

7.2.1 *Comparing an Individual's Threshold With a Literature Value*—The test may be conducted, for example, to diagnose anosmia or ageusia, or to study sensitivity to pain, noise, or odor. This is the simplest category requiring a minimum of 20 to 40 3-AFC presentations to the individual in question (see 2.2). A number of training sessions may be required to establish the range of concentrations that will be used and to make certain that the individual is fully familiar with the stimulus to be detected as well as the mechanics of the test.

7.2.2 *A Population Threshold is Required*, for example, the odor threshold of a population exposed to a given pollutant, or the flavor threshold of consumers of a beverage for a given contaminant. In this case, recourse must be had to the rules of sampling from a population (see Ref (1)⁵ and Practice E 122), which require the following:

- (1) That the population be accurately defined and delimited,
- (2) That the sample drawn be truly random, that is, that every member of the population has a known chance of being selected, and

(3) That knowledge of the degree of variation occurring within the population exists or can be acquired in the course of formulating the plan of sampling.

7.2.2.1 In practice, the cost and availability of panelists places serious limitations on the degree to which population factors affecting thresholds, for example, age groups, gender, ethnic origin, well versus ill, smoker versus nonsmoker, trained versus casual observers, etc., can be covered. The experimenter is typically limited to panels of 5 to 15, with each receiving 20 to 40 3-AFC presentations, for a total of 100 to 600 presentations. If the resulting thresholds are to have validity for the population, the experimenter should include the following steps:

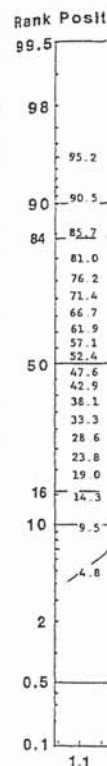
- (1) Calculate and tabulate the thresholds for each individual;
- (2) Repeat the test for those individuals (outliers) falling well beyond the range of the rest of the panel;
- (3) For any individuals whose threshold at first did not fall well within the range of samples presented to them, adjust the range and repeat the test; and
- (4) If needed to obtain a desired level of precision, repeat the test series with a second or third panel sampled from the same population of interest.

7.2.2.2 Thresholds vary with age, and one approach to a generalizable population value is to adjust thresholds obtained at various ages to an estimate for healthy 20-year-olds, using Amoore's finding (2) that between the ages of 20 and 65, odor threshold concentrations double for approximately each 22 years of age.

7.2.3 *The Distribution of Thresholds in the Population is Required*, for example, to determine what proportion of the population is affected by a given level of a pollutant, or, conversely, to determine which concentrations of a pollutant will affect a given percent of a population. The requirements for testing are the same as in 7.2.2, except that it is even more important to cover the range well, for example, to repeat the tests for those individuals whose thresholds fall in thinly populated parts of the panel range. Consideration should be given to increasing the number of presentations per concentration from 5-7 to 7-10 for such panel members. If the individual thresholds are plotted as in Fig. 1, any sector requiring study will be apparent from the graph.

7.3 *Trained Versus Casual Assessors*—Thresholds should normally be determined for assessors trained by repeated exposure to detect the stimulus in question whenever it is present; however, if the threshold sought is that of a casual observer (for example, for a warning agent in household gas), naive panelists and mild distraction (for example, noise) may be used (see Ref. (3)).

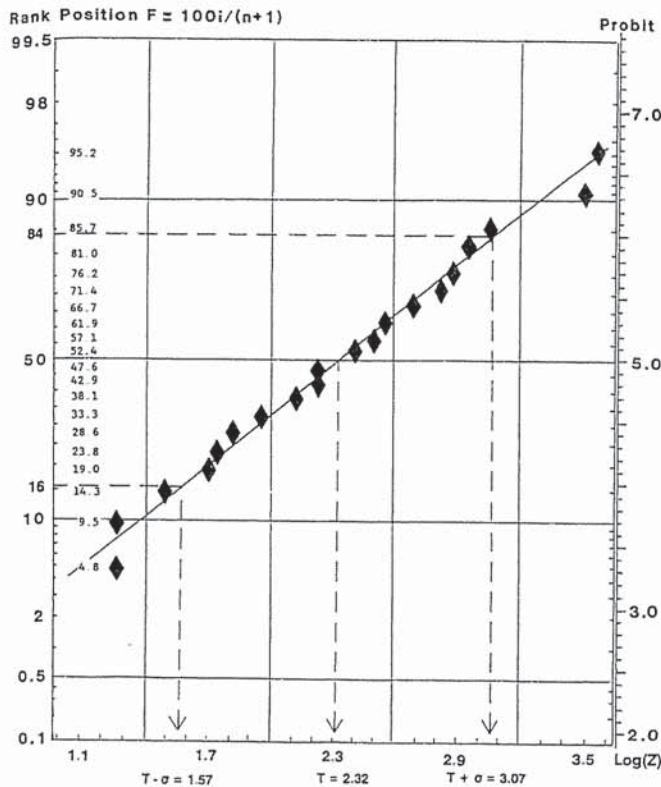
7.4 *Choice of the Measure of Central Tendency*—The report should contain a table or graph providing the individual thresholds of each observer. If a group threshold is required, the measure of central tendency chosen should be that which best represents the distribution obtained. In a few cases (Fig. 2), the results form a symmetrical, bell-shaped distribution, and the arithmetic mean, or median may be used. With sensory data, the distribution is typically skewed (Fig. 3); however, it may be normalized by converting the concentration units to log



NOTE 1—described in Group threshold variation from units. The $2.32 + (0.75 \text{ of the normal})$ FIG.

⁵ The boldface numbers in parentheses refer to the list of references at the end of this standard

E 1432 - 04



NOTE 1—This probability graph shows 20 panelists sorted by rank as described in 9.3.2. Data are adapted from French Standard X 43-101. Group threshold = $T = 50\%$ point = $\log(Z_{50}) = 2.32$. Group standard deviation from σ and 84 % points = $\sigma = (3.07 - 1.57)/2 = 0.75$ in $\log(Z)$ units. The 99 % point is off the graph but can be calculated as $2.32 + (0.75 \times 2.327) = 4.07$, where 2.327 is the % point on the abscissa of the normal curve of error.

FIG. 1 Group Threshold by Rank-Probability Graph

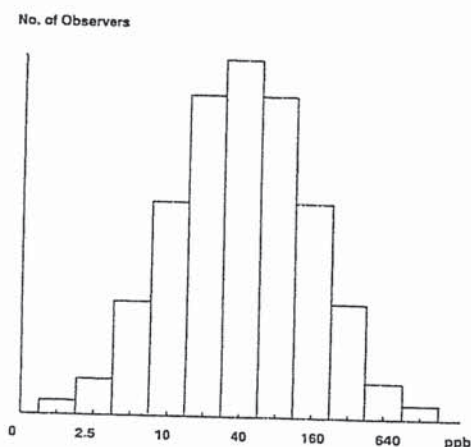


FIG. 2 Symmetrical, Bell-Shaped Distribution

concentration, which is equivalent to converting the arithmetic mean into the geometric mean. If, as is often the case, the distribution is irregular but approaches normality, the 50 % point of a log-probability graph (see Fig. 1) is the appropriate measure. Conversion of the concentration scale into double

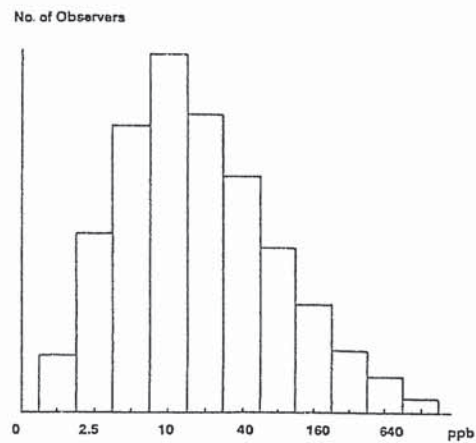


FIG. 3 Skewed Distribution

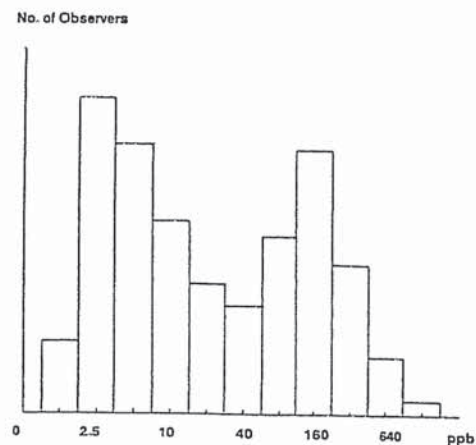


FIG. 4 Bi-Modal Distribution

logarithms (log of log) is occasionally needed to normalize a distribution. However, if the data show a bi-modal (Fig. 4) or multi-modal distribution, indicating the existence of sub-populations with different thresholds, the distribution cannot be normalized; instead, an attempt may be made to estimate the size and group threshold of each sub-population.

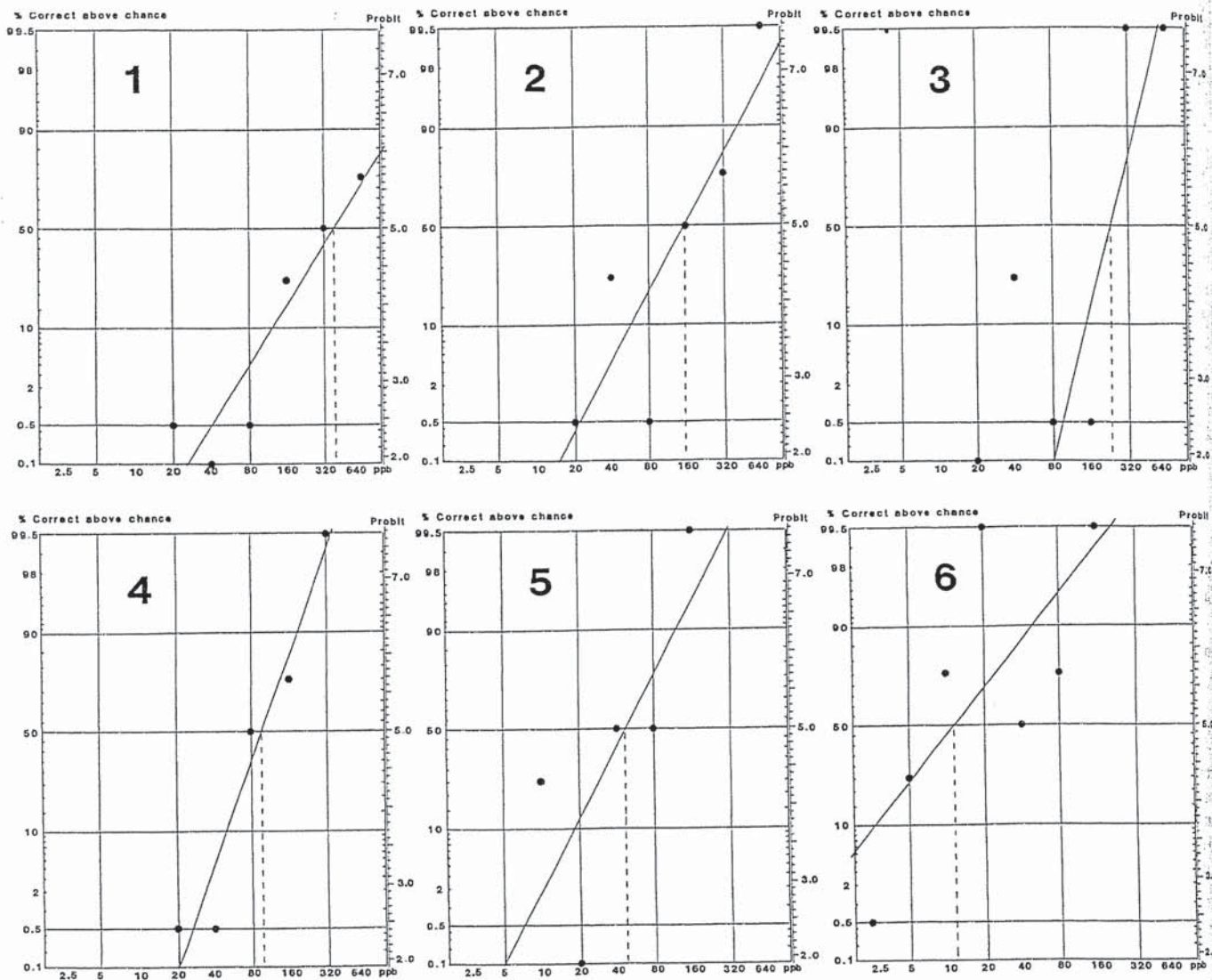
7.5 Group Standard Deviation—To characterize the dispersion of thresholds around the population mean, a group standard deviation, σ , may be estimated as shown in the examples, Figs. 5 and 6, and Fig. 1. This is permissible only if the distribution is normal or near-normal, or has been normalized.

8. Procedure

8.1 Tabulation and Transformation of Data—See Table 1.

8.1.1 Example 1: Threshold of Substance X in Purified Water—Six observers took part; each was tested at five or more concentrations chosen in advance⁶ to bracket each person's threshold; each took six tests per concentration, for a total of 30 to 36 presentations per observer (Table 1):

⁶ For example, by giving the person a single test (or a few tests) of the concentrations 640, 160, 40, 10, and 2.5 ppb.



NOTE 1—The results (using Probits and linear regression) are as follows:

Panelist No.	1	2	3	4	5	6
Threshold, ppb	381	166	226	97	47	12
Group standard deviation (six panelists), $\sigma = 0.539$ in log (ppb) units.						

FIG. 5 Graphic Estimation of Approximate Thresholds for the Six Panelists in 7.3

8.1.2 Convert Results to Percent Correct Above Chance, at each concentration level for each panelist, using the formula of the 3-AFC procedure:

$$\% \text{ correct above chance} = 100 \frac{\% \text{ correct} - \% \text{ correct by chance}}{100 - \% \text{ correct by chance}} = 100(3C - N)/2N \quad (2)$$

where:

N = number of tests presented per panelist and concentration (here, six), and

C = number of correct choices.⁷

8.2 Calculate the Threshold for Each Panelist:

8.2.1 Graphic Method—Plot percent correct choice above chance against log stimulus intensity on normal probability graph paper, as shown in Fig. 5. Plot scores of 100 % as 99.5 %, 0 % as 0.5 %, and less than 0 as 0.1 %; then fit a straight line through the points by eye. Read the threshold at that concentration corresponding to 50 % probability. Alternatively convert the percent scores to Probits (4) or use a table of the normal deviate. Fit the line by the method of least squares.

Paired-comparison and Duo-Trio
Triangular
Two-out-of-five

Correct above chance, % = $100(2C - N)/N$
Correct above chance, % = $100(3C - N)/2N$
Correct above chance, % = $100(10C - N)/9$

NOTE 1—iterative method NLIN procedure decreases ("panelists are

Method

Log (ppb)
Threshold, p
Group stand;

⁷ The formulas of other forced-choice procedures are:

E 1432 - 04

```

** PURPOSE: Fit logistic models  $P = (1/3 + \text{EXP}[K]) / (1 + \text{EXP}[K])$ ,
** where  $K = B(T - \text{LOG}[X])$ ,
**  $P$  is the proportion of correct identifications,
**  $B$  is the slope,
**  $X$  is the actual concentration (ppb) of Substance X
** in purified water,
** and  $T$  is the threshold value in  $\log(\text{ppb})$ .

```

```

PROC NLIN Method=DUD Data=Input; by panelist;
PARMS B=-4 T=2
K = B*(T - LOG10(K));
E = EXP(K);
N = (1/3 + E);
D = (1 + E);
MODEL P = N/D;
TITLE2 'Logistic Regression Models';

```

```

RUN;

```

```

OUTPUT FOR PANELIST 4:

```

```

Logistic Regression of Threshold Data Using SAS PROC NLIN
Logistic Regression Models
NON-LINEAR LEAST SQUARES ITERATIVE PHASE
DEPENDENT VARIABLE: P METHOD: DUD

```

ITERATION	B	T	RESIDUAL	SS
-3	-4	2.000000000	0.025885700365	
-2	-4.4	2.000000000	0.02054415598	
-1	-4	2.200000000	0.084958944779	
0	-4.4	2.000000000	0.02054415598	
1	-5.852958	1.961443385	0.010812277188	
2	-6.259745	1.967823308	0.010766524899	
3	-6.189164	1.951938036	0.010504941622	
4	-6.283542	1.954261395	0.010481402394	
5	-6.280162	1.954257276	0.010481361251	
6	-6.281544	1.954068199	0.010481219887	
7	-6.277816	1.953905805	0.010481193047	
8	-6.280506	1.953919346	0.010481176612	
9	-6.281737	1.953896400	0.010481176219	
10	-6.281715	1.953899496	0.010481176193	

```

CONVERGENCE CRITERION MET.

```

```

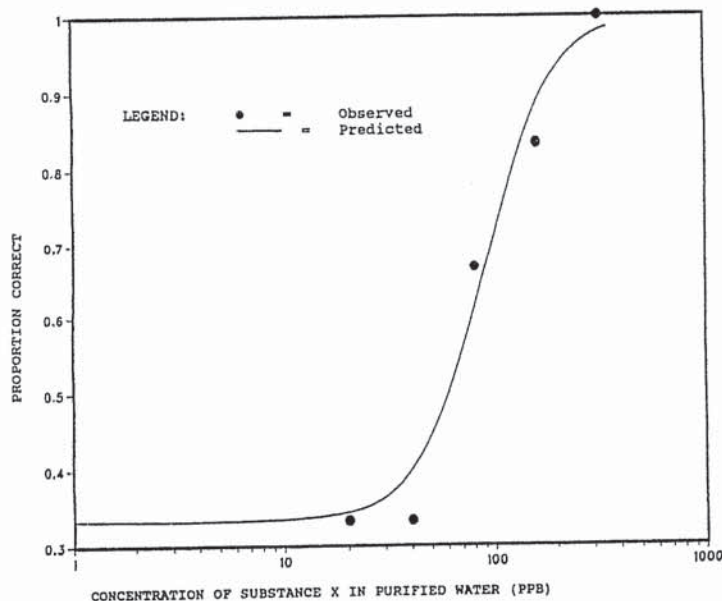
NON-LINEAR LEAST SQUARES SUMMARY STATISTICS DEPENDENT VARIABLE P
SOURCE DF SUM OF SQUARES MEAN SQUARE
REGRESSION 2 2.3500748238 1.1750374119
RESIDUAL 3 0.0104811762 0.0034937254
UNCORRECTED TOTAL 5 2.3605560000
(CORRECTED TOTAL) 4 0.3558448000

```

```

PARAMETER ESTIMATE ASYMPTOTIC STD. ERROR ASYMPTOTIC 95% CONFIDENCE INTERVAL LOWER UPPER
B -6.281714751 1.6824126163 -11.635992903 -0.9274366000
T 1.953899496 0.0473965533 1.803059965 2.1047390264

```



P	C	LOG[X]	X
.967	.95	2.4226	265
.933	.90	2.3037	201
.833	.75	2.1288	135
.667	.50	1.9539	90
.500	.25	1.7790	60
.400	.10	1.6041	40
.367	.05	1.4770	30

NOTE 1—The PROC NLIN fits nonlinear regression models by least squares. Following the regression expression, the operator selects one of four iterative methods (here, DUD) and must specify an approximate value for the parameters B (the slope, here = -4) and T (the threshold, here = 2). The NLIN procedure first prints out the starting values for B and T , then proceeds stepwise (here, ten steps) until the residual sum of squares no longer decreases ("convergence criterion met"). The threshold (here, $T = \log(\text{ppb}) = 1.954$) is found as the last value in the T column. The results for the six panelists are as follows:

Method	Panelist	1 DUD	2 DUD	3 DUD	4 DUD	5 DUD	6 MARQUARDT
Log (ppb)		2.518	2.249	2.368	1.954	1.806	0.892
Threshold, ppb		330	178	249	90	64	7.8
Group standard deviation (six panelists), $\sigma = 0.59$ in $\log(\text{ppb})$ units.							

FIG. 6 Output from SAS NLIN Program (6) with Details for Panelist No. 4

E 1432 - 04

TABLE 1 Number of Correct Responses for Each Panelist at Each Concentration^A

Concentrations presented, ppb	No. Correct					
	Panelist No.					
	1	2	3	4	5	6
640	5	6	6			
320	4	5	6	6		
160	3	4	2	5	6	6
80	2	2	2	4	4	5
40	1	3	3	2	4	4
20	2	2	1	2	0	6
10					3	5
5						3
2						2

Concentration presented, ppb	Correct Above Chance for Each Panelist at Each Concentration, % ^B					
	Panelist No.					
	1	2	3	4	5	6
640	75	100	100			
320	50	75	100	100		
160	25	50	0	75	100	100
80	0	0	0	50	50	75
40	-25	25	25	0	50	50
20	0	0	-25	0	-50	100
10					25	75
5						25
2						0

^AResults obtained in Example 1 (8.1.1).^BData converted per 8.1.2.

8.2.2 *Computer Package*—Use a computer package employing an iterative curve-fitting procedure and weighting the data by probability. The desired S-shaped curve (ogive) may be approximated using the normal probability curve or, alternatively, a logistic model (5):

$$P = (1/3 + e^k)/(1 + e^k) \quad (3)$$

$$k = b(t - \log[x])$$

where:

P = proportion of correct responses, that is, C/N ,

b = slope,

x = concentration (in ppb),⁸ and

t = threshold (in log(ppb) units).

Note that conversion per 8.1.2 is not used here. The threshold is at $P = 2/3$, and all values for C can be accommodated; also, $C = N$ and $C = 0$. Fig. 6 shows the results obtained for Panelist No. 4.

8.3 Group Threshold:

8.3.1 *Calculation of Group Threshold According to 7.4*—Report each individual threshold obtained. If the purpose of the test so requires, and the results themselves permit, a group threshold may be calculated according to 7.4. In Example 1 given in 8.1.1, the geometric mean may be chosen as the best central measure:

Panelist No.	Log T	Threshold, ppb
1	2.518	330
2	2.249	178
3	2.368	233

⁸ If x is on logarithmic form, for example, $x = \log(Z)$ as in Dynamic Triangle Olfactometry, the formula is $k = b(t - x)$, and t is obtained in $\log(Z)$ units.

4	1.954	90
5	1.806	64
6	0.892	7.8
sum	11.787	1.9645

Antilog = group threshold = group geometric mean = 92 ppb; group standard deviation (six panelists), $\sigma = 0.59 \log(\text{ppb})$ units.

8.3.2 *Group Threshold by Rank/Probability Graph*—Use this method (7) when the number of individuals is 10 to 15 or higher and the distribution is near normal. See the example in Fig. 1. Sort the panelist thresholds by rank i and plot them in a probability graph, using as ordinate one of the alternatives given in Ref (7) (there is no one accepted formula) or the

$$\text{"rank position"} F_i = 100 i/(n + 1) \quad (4)$$

8.3.2.1 For example, Panelist No. 11, out of a group of 20 will plot at $100 \times 11/(20 + 1) = 52.4\%$. If, as here, a straight line can be drawn through the points, consider the group normally distributed with group threshold at the 50 % point and group standard deviation (one sigma) limits at the 16 % and 84 % points. Read other points of interest from the graph; for example, read the concentration that only 1 % of the population can detect as the 99 % point or, conversely, find the which 95 % can detect as the 5 % point.

9. Presentation of Results

9.1 Report all test conditions, such as the nature and source of the samples, method of sampling, choice of control sample (diluent), equipment and physical test setup under which samples were presented to the panelists, concentrations or flowrates used, temperature and other conditions of the samples, and instructions and report sheets given to the panelists.

9.2 Report the composition of the panel with regard to age, gender, and experience. Additional information may be useful, for example, familiarity with the stimulus being evaluated, health, smoking, use of dentures, time since last meal, etc. No panelist should be identified by name; nor should the report allow a reader familiar with the panel to refer a particular judgment to a particular panel member.

9.3 Report the number of repetitions of the presentations per panelist.

9.4 Report the individual thresholds and, if the purpose of the test so requires and the results themselves so permit, calculate a group threshold and a group standard deviation, as shown in Figs. 5 and 6, and Fig. 1.

10. Precision and Bias

10.1 Because sensory threshold values are functions of sample presentation variables and of individual sensitivities, interlaboratory tests cannot be interpreted statistically in the usual way, and a general statement regarding precision and bias of thresholds obtained by this practice cannot be made. However, certain comparisons made under particular circumstances are of interest and thus are detailed below.

10.2 When four panels of 23 to 35 members evaluated butanol in air (8) in the same laboratory, the ratio of the highest to the lowest panel threshold was 2.7 to 1; when the same panel repeated the determination on four days, the ratio was 2.4 to 1. For ten panels of nine members evaluating hexylamine in air

the rat
Practic
10.3
purifie
highes
Interla
thyl ac
somew
10.4
with se

(1) Snee
lowe
(2) Amc
(3) Amc
Odo
ties f
of A
(4) Finn
1971
(5) Bish
Anal
(6) SAS
NC,



E 1432 - 04

the ratio was 2.1 to 1. Although the method used was that of Practice E 679, the results are comparable.

10.3 When 14 laboratories determined the threshold of purified hydrogen sulfide in odorless air (9), the ratio of the highest to the lowest laboratory threshold was 20 to 1. Interlaboratory tests with dibutylamine, isoamyl alcohol, methyl acrylate, and a spray thinner for automobile paint gave somewhat lower ratios.

10.4 An extreme form of bias is lack of experience, either with sensory testing in general or with the substance under test.

In several trial series with vanillin in aqueous solution (10), untrained panels reported thresholds up to 1000-fold higher than trained panels.

11. Keywords

11.1 air pollution; odor; panel; sensory evaluation; taste; 3-Alternative Forced-Choice Presentation; threshold; water pollution

REFERENCES

- (1) Snedecor, G. W., and Cochran, W. G., *Statistical Methods*, 7th ed., Iowa State University Press, Ames, IA, 1980, Chapter 17.
- (2) Amoores, J. E., Personal Communication to Task Group E18.04.25.
- (3) Amoores, J. E., and Hautala, E., "Odor as an Aid to Chemical Safety: Odor Thresholds Compared with Threshold Limit Values and Volatilities for 214 Industrial Chemicals in Air and Water Dilution," *Journal of Applied Toxicology*, Vol 3, No. 6, 1983, pp. 272-290.
- (4) Finney, D. J., *Probit Analysis*, 3rd ed., Cambridge University Press, 1971.
- (5) Bishop, Y., Fienberg, S., and Holland, P., *Discrete Multivariate Analysis*, MIT Press, Cambridge, MA, 1980, pp. 357-358.
- (6) SAS User's Guide: Statistics, Version 5 Edition, SAS Institute, Cary, NC, 1985, pp. 575-606.
- (7) Snedecor, G. W., and Cochran, W. G., *Statistical Methods*, 7th ed., Iowa State University Press, Ames, IA, 1980, Chapter 4, pp. 59-63.
- (8) Dravnieks, A., Schmidtsdorff, W., and Meilgaard, M., "Odor Thresholds by Forced-Choice Dynamic Triangle Olfactometry: Reproducibility and Methods of Calculation," *Journal of the Air Pollution Control Association*, Vol 36, 1986, pp. 900-905.
- (9) German Standard VDI 3881, Part 1, *Olfactometry. Odour Threshold Determination. Fundamentals*, Verein Deutscher Ingenieure, VDI-Verlag GmbH, Düsseldorf 1986, pp. 25-27.
- (10) Powers, J. J., and Shinholser, K., "Flavor Thresholds for Vanillin and Predictions of Higher or Lower Thresholds," *Journal of Sensory Studies*, Vol 3, 1988, pp. 49-61.

ASTM International takes no position respecting the validity of any patent rights asserted in connection with any item mentioned in this standard. Users of this standard are expressly advised that determination of the validity of any such patent rights, and the risk of infringement of such rights, are entirely their own responsibility.

This standard is subject to revision at any time by the responsible technical committee and must be reviewed every five years and if not revised, either reapproved or withdrawn. Your comments are invited either for revision of this standard or for additional standards and should be addressed to ASTM International Headquarters. Your comments will receive careful consideration at a meeting of the responsible technical committee, which you may attend. If you feel that your comments have not received a fair hearing you should make your views known to the ASTM Committee on Standards, at the address shown below.

This standard is copyrighted by ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959, United States. Individual reprints (single or multiple copies) of this standard may be obtained by contacting ASTM at the above address or at 610-832-9585 (phone), 610-832-9555 (fax), or service@astm.org (e-mail); or through the ASTM website (www.astm.org).